journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM

Azam Davari Dolatabadi [a], *Siamak Esmael Zadeh Khadem* [a,*], *Babak Mohammadzadeh Asl* [b]

[a] *Mechanical Engineering Department, Tarbiat Modares University, Tehran, Iran*
[b] *Electrical and Computer Engineering Department, Tarbiat Modares University, Tehran, Iran*

ARTICLE INFO

Article history:
Received 25 June 2016
Received in revised form
7 September 2016
Accepted 18 October 2016

Keywords:
Electrocardiogram
Coronary artery disease
Hearth rate variability
Principal component analysis
Support vector machines

ABSTRACT

Background and objective: Currently Coronary Artery Disease (CAD) is one of the most prevalent diseases, and also can lead to death, disability and economic loss in patients who suffer from cardiovascular disease. Diagnostic procedures of this disease by medical teams are typically invasive, although they do not satisfy the required accuracy.
Methods: In this study, we have proposed a methodology for the automatic diagnosis of normal and Coronary Artery Disease conditions using Heart Rate Variability (HRV) signal extracted from electrocardiogram (ECG). The features are extracted from HRV signal in time, frequency and nonlinear domains. The Principal Component Analysis (PCA) is applied to reduce the dimension of the extracted features in order to reduce computational complexity and to reveal the hidden information underlaid in the data. Finally, Support Vector Machine (SVM) classifier has been utilized to classify two classes of data using the extracted distinguishing features. In this paper, parameters of the SVM have been optimized in order to improve the accuracy.
Results: Provided reports in this paper indicate that the detection of CAD class from normal class using the proposed algorithm was performed with accuracy of 99.2%, sensitivity of 98.43%, and specificity of 100%.
Conclusions: This study has shown that methods which are based on the feature extraction of the biomedical signals are an appropriate approach to predict the health situation of the patients.

1. Introduction

Cardiovascular disease takes the lives of about 17 million people each year worldwide. Coronary Artery Disease (CAD) has reached nearly epidemic proportions in most of the societies and is the cause of more deaths than any other group of cardiovascular diseases [1]. CAD is a condition where the arteries which supply blood to the heart muscles are hardened and narrowed resulting in buildup of plaques. This generally occurs due to growth of plaque within the arteries which contributes to the reduction of the amount of blood flows and also the amount of oxygen [2]. This malfunction can weaken the heart muscles and ultimately may lead to heart failure [3]. Therefore, early stage detection of CAD has a great deal of importance.

There are several diagnostic methods for CAD which generally begins with the onset of one of the common symptoms of the disease like heart attack or a sudden cardiac arrest. Some

[*] *Corresponding author.* Mechanical Engineering Department, Tarbiat Modares University, Tehran, Iran.
E-mail addresses: azam.davari@modares.ac.ir (A. Davari Dolatabadi); khadem@modares.ac.ir (S.E.Z. Khadem).

http://dx.doi.org/10.1016/j.cmpb.2016.10.011
0169-2607/© 2016 Elsevier Ireland Ltd. All rights reserved.

of the general diagnostic tests include stress test, Electrocardiography (ECG), echocardiography, coronary angiography or cardiac catheterization. Symptoms of CAD are generally diagnosed through a treadmill stress test which is painful to the patients and causes discomfort to them. Though ECG is used during several analyses, one of the major limitations of this technique is the invisible symptoms of CAD in an ECG record [4]. Hence, the only alternative is the coronary angiography or cardiac catheterization which is an invasive methodology and procedures are specialized processes which involve a lot of time, effort and cost. In addition, these can only be conducted by trained people [3]. These limitations can be eliminated easily by using computer aided system for the diagnosis of heart disorders.

Digital signal processing techniques are helpful and non-invasive tools used in the diagnosis of CAD. These techniques can be performed on both ECG and HRV, but HRV includes nonlinear and non-stationary nature of heart activity. Actually it is computed from variation over time of the period between consecutive heartbeats (RR intervals) [5]. This signal has been widely used as an appropriate tool for assessing the situation of the heart [6–12]. Linear and non-linear features of HRV have been evaluated for detection of cardiovascular disease.

In 1994 linear analysis of HRV showed that the circadian rhythm decreases in CAD subjects [13]. It has been proved the correlation between CAD severity and a reduction of low-frequency power and decrease in high frequency power. In the time and frequency domain analysis, it has shown that the features are lower for CAD subjects. But in further research, it has been proved that because of stationary nature of heart signal the features vary over time and hence time domain analysis is not helpful enough for detecting CAD subjects from normal subjects. It must have taken into account that as the noise in signal increases, the effectiveness of frequency domain analysis will decrease. So it became clear that for this type of malfunction nonlinear features are more helpful [14].

Computer aided diagnostic methods which extract relevant features and use them in classifiers for automated detection of diseases can overcome these difficulties. Such techniques are noninvasive and provide reproducible and objective diagnoses, and hence can prove to be valuable adjunct tools in clinical practices [13].

The number of features which are extracted for early detection of CAD in this research has to be decreased in order to increase the accuracy of the prediction. Various statistical techniques are being effectively used in order to diminish the size of features. In this study in which the major focus is on the early detection, it is very important to not lose even small changes in the heart signal. So choosing the algorithm for dimension reduction has an important role in the accuracy of the detection. According to the results provided by Babaoğlu, PCA reduced dataset diagnostic accuracy rate, and decreases the sum of the training and test time and also training error in the determination of CAD [15].

SVM, proposed by Vapnik et al. [16,17], has recently gained wild attention in pattern recognition application fields, such as face recognition and detection [18,19]. The learning strategy of SVM is based on the principle of structural risk minimization (SRM), which makes SVM receive much attention in the past few years. The principle of SRM minimizes not only the observed risk (training errors) but also the simplification miscalculation bound during the training. On the other hand, the learning strategy of traditional learning machines, such as multilayer neural networks, is only formulated based on the principle of observed risk minimization. It has been shown that SVM is superior to other traditional learning machines [20] such as BPN, since SVM is able to gain better globalization ability for unseen data. It is known that how to obtain high classification accuracy for unseen HRV patterns based on few available HRV training patterns is the key issue in studying the problem of HRV classification. This issue dominates the reliability of a CAD recognition system. For HRV classification, SVM can be a promising approach. Therefore, in this paper, SVM is introduced into features for gaining high generalization ability in HRV classification.

The objective of our study is automatic diagnosis of CAD by conducting non-invasive tools on the HRV. In order to deal with the nonlinear nature of heart signal, HRV has been extracted from ECG and nonlinear features like fuzzy entropy are extracted. By using PCA algorithm it has been shown that frequency and time domain features are not appropriate indicators for distinguishing CAD from normal subjects. The contribution of this work is to extract 11 nonlinear features which are more effective for early detection. Actually nonlinear features are exhibiting nonlinear nature of HRV, but it has been shown that some of the morphological changes due to coronary artery stenosis, which is a stage before CAD disease, are detectable through some of the nonlinear features. We also propose an optimized SVM in order to automate classification of normal and CAD data using these nonlinear features. The proposed classification has reached the accuracy of 99.2% by using 11 features which are the essential ones.

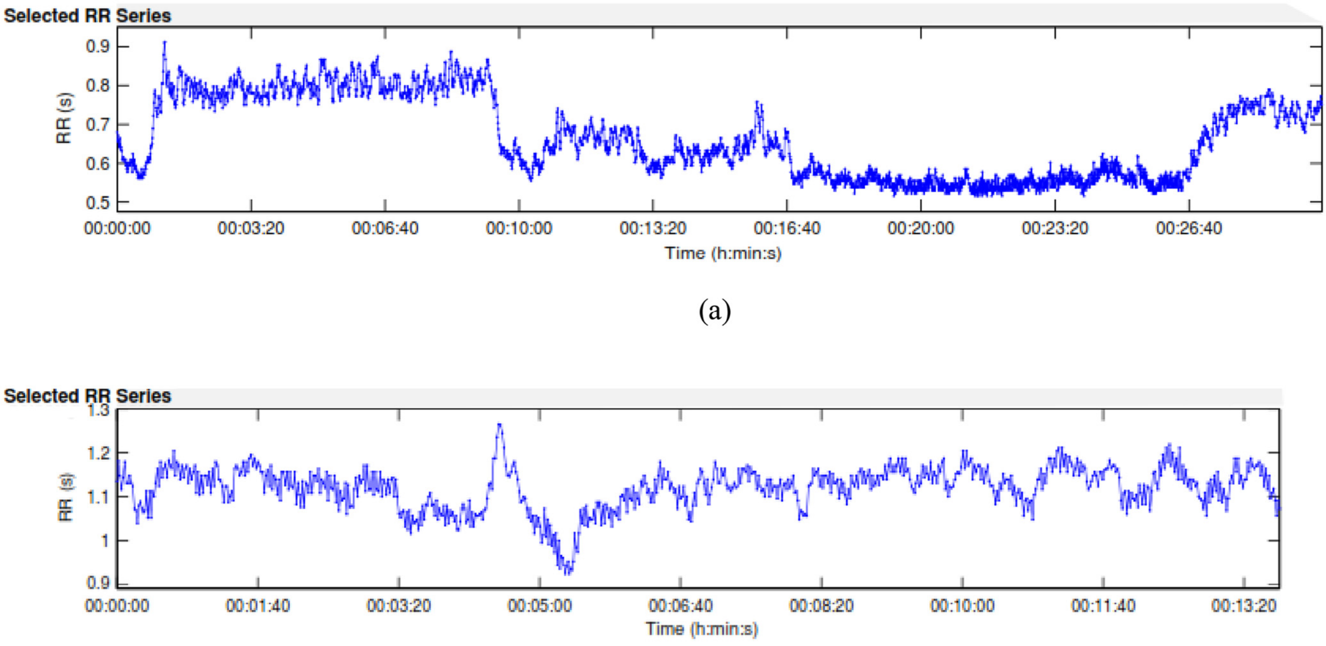
Our proposed CAD diagnosis method is completely automated. It contains two major steps: (i) training and (ii) testing. During training stage, the nonlinear features are used in training the classifier. In further stage, which is called testing, the same nonlinear features are extracted and served as an input for the pre-trained classifier for automatic detection. This method can be a helpful assistant for clinician in order to eliminate human errors and risk of invasive detections.

2. Data acquisition

2.1. *Data set introduction*

The data for the CAD group were obtained from 86 lengthy ECG recordings of 80 human subjects. The subjects were 46 men, aged 44–85 years, and 29 women, aged 23–87 years. These recordings were from the Long-Term ST Database [21]. The Long-Term ST Database contains 86 lengthy ECG recordings of 80 human subjects, chosen to exhibit a variety of events of ST segment changes, including ischemic ST episodes, axis-related non-ischemic ST episodes, episodes of slow ST level drift, and episodes containing mixtures of these phenomena [21]. In the case that the study is about coronary artery disease, 23 subjects from this database who only suffer from CAD have been chosen.

The data for the normal control group were obtained from 24-hour Holter monitor recordings of 54 healthy subjects (30



(a)

(b)

Fig. 1 – RR intervals signal of: (a) normal and (b) CAD.

men and 24 women, aged 29–76 years with mean of 61). These recordings were from the Normal Sinus Rhythm RR Interval Database [22]. In order to have an equal size of database, 23 records of this database have been chosen.

2.2. *Pre-processing*

The ECG signals were passed via a low-pass filter with a cut-off frequency of 20 Hz to eliminate low frequency noise. By passing through a high-pass filter with a lower cut-off frequency of 0.3 Hz, the baseline wander was also removed. Power source interference noise was removed using a 50 Hz notch filter. Finally, the R peaks of ECG signal were detected using the Pan-Tompkins algorithm [23] to accurately extract the QRS complexes. This is a real-time algorithm that employs a band-pass filter, differentiator, and integrator over a moving window. The RR interval was computed as the interval between two successive QRS complexes (R peaks). Heart rate (beats per minute) is calculated from the RR interval (in seconds) using:

$$HR_{bpm} = \frac{60}{RR} \tag{1}$$

The typical heart rate variability signals of a normal subject and a subject with CAD are shown in Fig. 1.

3. Feature extraction

In this section, time, frequency and nonlinear domain features which were used for diagnosing CAD patients from normal subjects have been discussed.

3.1. *Time domain features*

Most of the time domain features are derived from the RR intervals. The clearest features are mean RR and mean HR. Some of them describe variability within the RR intervals such as SDNN, SDSD, RMSSD, and PNN50 which are statistical, while others such as TINN and HRV triangular index are geometric features. The definition of these features is listed in Table 1.

Analysis of the data in time domain has been performed using mentioned features. The results of this analysis and the difference of their values have been listed in Table 2 for CAD and normal cases. Regarding the p-value it has been shown that the time domain features for this data set are not a proper tool for diagnosis of the CAD.

Table 1 – The definition of the time domain features.

Features	Description
SDNN	Standard deviation of normal to normal R-R intervals
SDSD	The standard deviation of successive RR interval differences
RMSSD	Square root of the mean of the sum of the squares of differences between adjacent NN intervals
PNN50	Square root of the mean of the sum of the squares of differences between adjacent NN intervals
TINN	The baseline width of the RR histogram evaluated through triangular interpolation
HRV triangular index	Number of all NN intervals/maximum number

Table 2 – Mean value of time domain features for normal and subjects and p-value of each feature.

Features	Normal	CAD	p-value
SDNN	808.15	698.67	0.136
SDSD	51.9	45.4	0.215
RMSSD	142.11	104.4	0.143
PNN50	28.6	24.2	0.124
TINN	569.8	618.3	0.142
HRV triangular index	16.1	13.4	0.117

3.2. *Frequency domain features*

In the frequency domain method, a power spectrum density (PSD) estimate is calculated for the RR interval series. The regular PSD estimators implicitly assume equidistant sampling; thus, the RR interval series are converted to equidistantly sampled series by interpolation methods prior to PSD estimation. In HRV analysis, the PSD estimation is generally carried out using either FFT based or parametric AR modeling based methods. Details for these methods have been provided in Reference [24]. The advantage of FFT based methods is the simplicity of implementation, while the AR spectrum yields improved resolution especially for short length samples. Another property of AR spectrum which has made it popular in HRV analysis is its capability to be factorized into separate spectral components. The disadvantages of the AR spectrum, however, are the complexity of model order selection and the contingency of negative components in the spectral factorization. As a result, it may be advantageous to calculate the spectrum with both methods to have comparable results [25].

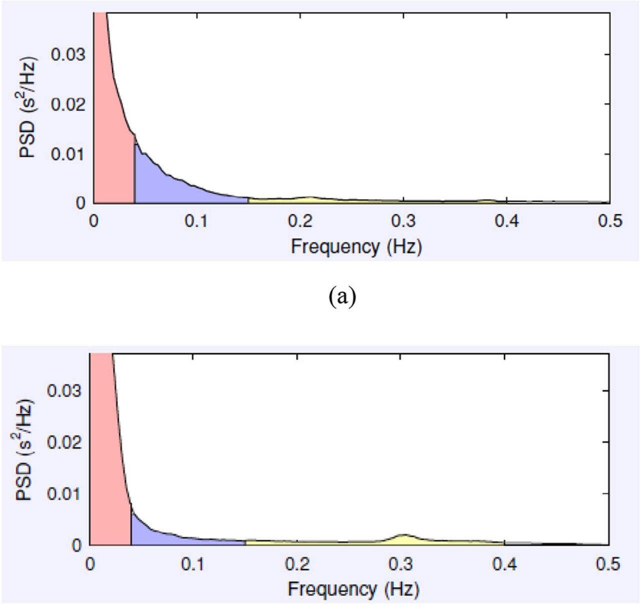
There are three main frequency regions in the heart rate signal:

- The power in the frequency range of 0.15 Hz–0.5 Hz is defined as a high frequency (HF) power band.
- The power in the frequency range of 0.0.4 Hz–0.15 Hz is defined as a low frequency (LF) power band.
- The power in the frequency range of 0.0033 Hz–0.04 Hz is defined as a very-low-frequency (VLF) power band.

HF region is an indicator of the vagal activity and respiratory sinus arrhythmia (RSA) while LF refers to the baroreceptor control mechanisms and the combined effect of sympathetic and vagal systems. The VLF power spectrum is an indicator for the vascular mechanisms and rennin-angiogenesis systems. In this study, total power, HF and LF were measured as well as the ratio of LF power to HF power [13]. Typical PSD plots for a normal subject and a subject with CAD are shown in Fig. 2.

3.3. *Nonlinear features*

Because of the nonlinear nature of biological signals, the theory of nonlinear dynamics is frequently used to analyze them. The following are the nonlinear methods that are used in this study: Poincare plots, Recurrence Quantification Analysis (RQA) parameters, Approximate Entropy (ApEn), Sample Entropy (SampEn), Detrended Fluctuation Analysis (DFA), and Correlation Dimension (D_2).



(a)

(b)

Fig. 2 – Typical PSD of heart rate signal: (a) normal, (b) CAD.

3.3.1. *Poincare plots*

Poincare plot is a visual plot, which was adopted from nonlinear methods to study the behavior of RR interval variability. It depicts the correlation between consecutive intervals in graphical representation [13]. The short term variability (SD1) of the heart signal is measured by the points that are perpendicular to the line-of-identity, and long term variability (SD2) by the points along the line-of-identity [26]. By looking at the Poincare plot shapes, normal subjects can be discriminated from subjects with CAD. For subjects with CAD, SD2 and SD1 are very low compared to the normal ones. Poincare plots for a normal subject and a subject with CAD have been shown in Fig. 3.

3.3.2. *Recurrence quantification analysis (RQA)*

Recurrence plot (RP) indicates the times at which path of the phase space meets the same location in the phase space for a given instant of time. The duration and counts of recurrences of the dynamical systems are estimated by RQA. It measures the dynamicity and subtle rhythmicity in the HR signal. The RQA parameters evaluate the complexity and non-stationary nature of the time series. Zbilut et al. [27] showed the usefulness of RQA in detecting randomness and complexity in non-stationary heart beats which cannot be analyzed easily by conventional techniques [13].

3.3.3. *Approximate entropy (ApEn)*

Approximate entropy (ApEn) measures the complexity or irregularity of the signal [28,29]. High disordered signal indicates a large value of ApEn. ApEn can be formulated as:

$$ApEn(m, r, N) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\log C_i^m(r) - \frac{1}{N-m}\sum_{i=1}^{N-m}\log C_i^{m+1}(r) \quad (2)$$

where C_m^i is the correlation integral and expressed as:

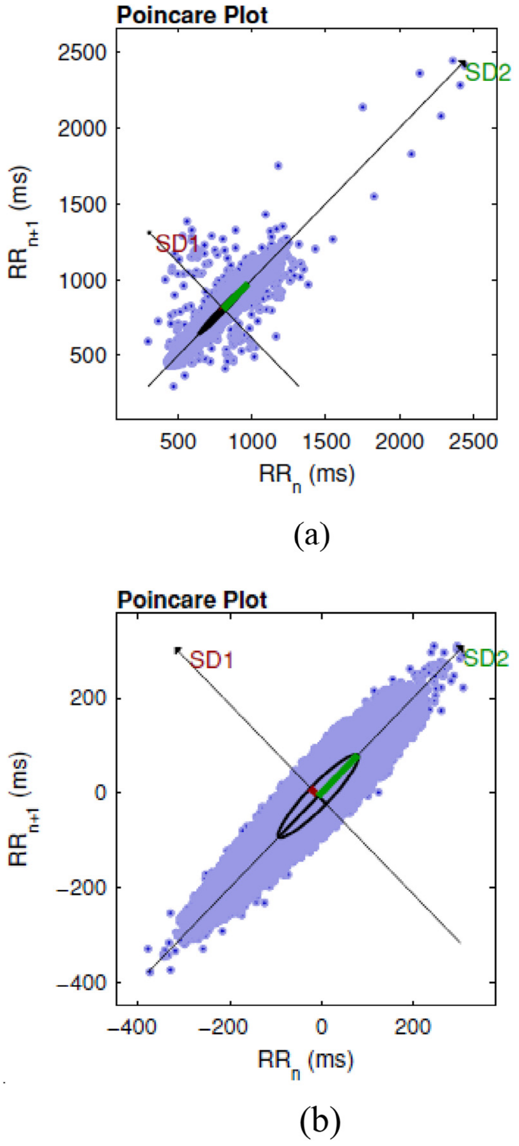


Fig. 3 – Typical Poincare plots of HR signals: (a) normal, (b) CAD.

$$C_i^m(r)=\frac{1}{N-m+1}\sum_{j=1}^{N-m+1}\Theta(r-\|x_i-x_j\|) \tag{3}$$

where x_i, x_jstand for phase trajectory points and N, r, Θ indicate number of points in the phase space, radial length of a circular disk centered at the reference points, step function and embedding dimension respectively.

In this study, the default value of *m* is set to 2. The length of the data (N) can also affect ApEn. There is a relation between tolerance r and standard deviation of the data (SDNN). This relation provides a way to compare different data types. A common selection for *r* is 0.2 × SDNN, which is also the default value in this work [30].

3.3.4. *Sample entropy (SampEn)*

SampEn is a quantifier for the complexity in a signal. Higher values of SampEn describe more irregularities in the time series. It is more refined than ApEn. In order to evaluate sample entropy, continuous matching of points inside the radius r is done as long as match exists. The variables *A(k)* and *B(k)* for all lengths *k* up to *e* keep track of all matching templates. It is given by:

$$SampEn(k,r,N)=-\ln\frac{A(k)}{B(k-1)} \tag{4}$$

In this study, for *k* = 0, 1, ..., *m* − 1 with *B*(0) = *N*, the length of the HR signal, *r* was taken equal to 0.2 and *m* maximum template length was set to 2 [13].

3.3.5. *Detrended fluctuation analysis*

Detrended Fluctuation Analysis assesses the self-similar properties of short term HR signals [31]. It also measures the correlation within the signal for different time scales [32]. The RR interval time series are integrated, and then divided into segments of equal lengths n, and a least squares line $y_n(k)$ is fitted into the data within each segment. Next the integrated series y(k) is detrended by subtracting the local trend within each segment and the root-mean-square fluctuation of this integrated and detrended time series is calculated by:

$$F(n)=\sqrt{\frac{1}{N}\sum_{K=1}^{N}(y(k)-y_n(k))^2} \tag{5}$$

A linear relationship on a double log graph indicates presence of fractal scaling and the fluctuations can be characterized by scaling exponent α (the slope of the regression line relating log F(n) to log n) [25]. This graph for two different classes is shown in Fig. 4.

3.3.6. *Correlation dimension (D2)*

D2 is a useful measure of self-similarity of a signal [33]. Considering the algorithm [33], Correlation integral (C(r)) function is computed using the distances between a pair of points described by $s(i,j)=|x_i-x_j|$. C(r) is given by:

$$C(r)=\frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1,x\neq y}^{N}\Theta(r-|x_x-x_y|) \tag{6}$$

Correlation dimension (D2) can be estimated by:

$$D2=\lim_{r\to 0}\frac{\log[C(r)]}{\log(r)} \tag{7}$$

More variations in RR will lead to higher values of D2 and vice versa.

4. Introduction of utilized algorithms

4.1. *Principal component analysis (PCA)*

Principal component analysis (PCA) is one of the most valuable analyses from applied linear algebra. PCA is used abundantly in all forms of analysis and varies from neuroscience to computer graphics because it is a simple,

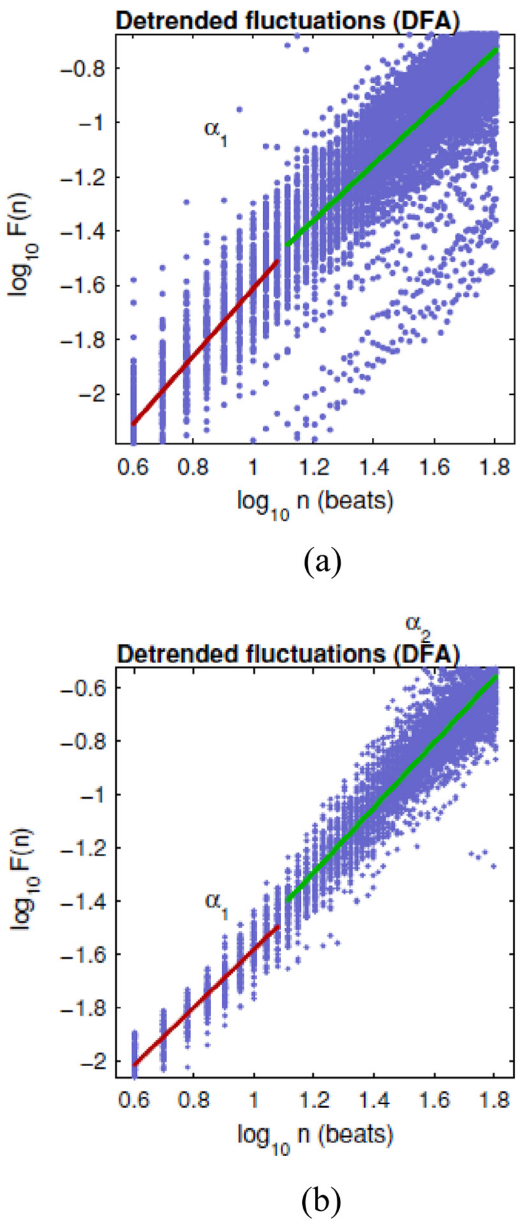


Fig. 4 – Typical DFA plot of HR signals: (a) normal, (b) CAD.

non-parametric method of extracting relevant information from confusing data sets. With minimal additional effort, PCA provides a roadmap on how to reduce a complex data set to a lower dimension to reveal the sometimes latent, simplified dynamics that often underlie it [14].

Consider given data points:

$$x_1, x_2, \ldots, x_n$$

We aim at reconstruction of these data with the dimension of D in a space with dimension M in which $M < D$, and the variance of the data will be maximized.

The variance of the mapped data will be defined by the equation, in which u_1 is the direction vector of the mapped space, N is the number of data, $\bar{x}$ is the average of the $\mathbf{x}_n$, S is the covariance matrix of the data and each x_n will be mapped to $u_1^T x_n$ which have to be maximized with regard to u_1.

$$\frac{1}{N}\sum_{n=1}^{N}\{u_1^T x_n - u_1^T \bar{x}\}^2 = u_1^T S u_1$$

Derivative with respect to (u_1) results in variance that is given with the equation:

$$u_1^T S u_1 = \lambda_1$$

Consequently when the eigenvector is equal to u_1, the variance will be maximized. This vector is called Principal Component.

The basis of this algorithm is elimination of the common parts of the problem. For this purpose, it is necessary to compute eigenvalues and eigenvectors of the data covariance matrix, and sort eigenvectors in the descending order of eigenvalues, and finally mapping the actual data into the directions of sorted eigenvectors.

4.2. *Support vector machine (SVM)*

The SVM is a single layer and highly nonlinear network based on statistical learning principles. It has ability to classify latent or not clearly observable patterns correctly [34–37]. Unlike other classifiers, SVM minimizes the structural risk rather than empirical risk. During the training of SVM, it maximizes the distance from patterns to the class separating hyper-plane. Generally, the patterns are not linearly separable, therefore nonlinear kernel transformation is performed [13]. Several kinds of kernel functions can be used by SVM such as: quadratic, polynomial and Radial Basis Function “RBF” kernels. In this study, the RBF kernel is used for the SVM.

Considering RBF kernel function for SVM, two important parameters should be considered. The first one is the cost parameter (C) which controls over fitting of the model. The next parameter that has to be taken into account is sigma (r) which controls the degree of nonlinearity of the model. Actually, the effectiveness of SVM depends on the selection of kernel's parameters and the best combination of them. In this paper, each combination of parameter choices is checked using cross validation, and the parameters with best cross-validation accuracy are picked. So, with test and trial process the optimization of this parameter is performed. The primal optimization problem for SVM is as follows:

$$min_w \quad f(w)$$

$$\text{s.t.} \quad g_i(w) < 0, \quad \mathrm{i} = 1, \ldots, \mathrm{k}$$

$$h_i(w) = 0, \quad i = 1, \ldots, l$$

The dual problem is

$$\underbrace{max}_{\alpha,\beta:\alpha_i\geq 0} \theta_D(\alpha,\beta) = \underbrace{max}_{\alpha,\beta:\alpha_i\geq 0} \underbrace{min}_{\omega} \pounds(\alpha,\beta,\omega)$$

It has been shown in this paper that this optimized SVM has led to more accuracy and sensitivity. The final SVM parameters obtained are: cost constant C: 3.39, and sigma: 5.91.

Table 3 – Principal component for CAD and normal classes.

Features	CAD (mean ± SD)	Normal (mean ± SD)
PC1	1.035847 ± 0.422798	1.147529 ± 1.376741
PC2	1.215651 ± 0.463969	1.639498 ± 0.651542
PC3	−0.40301 ± 0.220686	−0.91964 ± 0.163242
PC4	−0.42791 ± 0.131389	−0.54532 ± 0.288881
PC5	−0.62086 ± 0.230752	−0.59782 ± 0.127751
PC6	1.212541 ± 0.160244	1.19393 ± 0.16431
PC7	0.004219 ± 0.12498	0.022751 ± 0.129049
PC8	−0.26016 ± 0.102448	−0.28921 ± 0.091421
PC9	1.096657 ± 0.098897	1.131937 ± 0.061401
PC10	−0.44676 ± 0.082837	−0.4299 ± 0.069697
PC11	0.183839 ± 0.067236	0.174859 ± 0.033618

Table 4 – Results of classification using SVM.

TN	FN	TP	FP	Acc	Sn	Sp
128	2	126	0	99.2%	98.43%	100%

5. Results

In this study, ECG signal of 23 normal subjects database and 23 of CAD subjects was used as the dataset. All of them were analyzed in time domain and frequency domain. Because of the nonlinear nature of biological signals, nonlinear approach was used to determine nonlinear features of ECG. There were 9 features in time domain, 24 features in frequency domain (both FFT and AR methods) and 12 features in nonlinear domain. More dimensions will lead to more time, cost and redundant complexity; moreover, some features are not distinct in subjects with CAD and normal subjects. So, indiscriminative features have been excluded in order to reduce the dimension and computational complexities of the problem using PCA method. 45 features extracted from mentioned methods were used as inputs for PCA input. In this dimensionality reduction method (PCA), the dimensionality is reduced from 45 coefficients to 11 coefficients. These 11 features are classified using SVM classifier to one dimensional binary space, where 0 indicates normal subjects and 1 indicates subjects with CAD. Table 3 provides the mean and Standard Deviation (SD) of the first 11 clinically significant Principal Components (PC)—features extracted from PCA—for normal and CAD groups.

Data have been trained in an algorithm based on the pattern which has been drawn from the feature. This algorithm has been used for classifying the test data. Performance of this algorithm has been shown in Fig. 5.

The eleven features extracted through this technique were used to test the SVM classifier using cross validation. We have performed cross validation using 46 normal and CAD data signals. In each fold there were 16 data files of each class (normal and CAD), i.e., 2/3 of total number of data in each fold are taken as Train and 1/3 as Test data in every step. This pattern was repeated until the entire space was covered, i.e. the process is repeated 16 times till every fold can be taken as the test data. Performance measures were sensitivity, specificity and accuracy which were calculated on test data.

The independent samples (or two-sample) t-test is used to compare the means of two PCA classes. The results show that features that are extracted for two case studies are significantly different. The p-value less than .001 shows that discrimination between CAD and Normal subjects is true.

The number of True Positives (TP), False Negatives (FN), True Negatives (TN), and False Positives (FP) obtained using SVM classifier and PCA dimensionality reduction technique is provided in Table 4, where the concept of TN (True Negative) is the

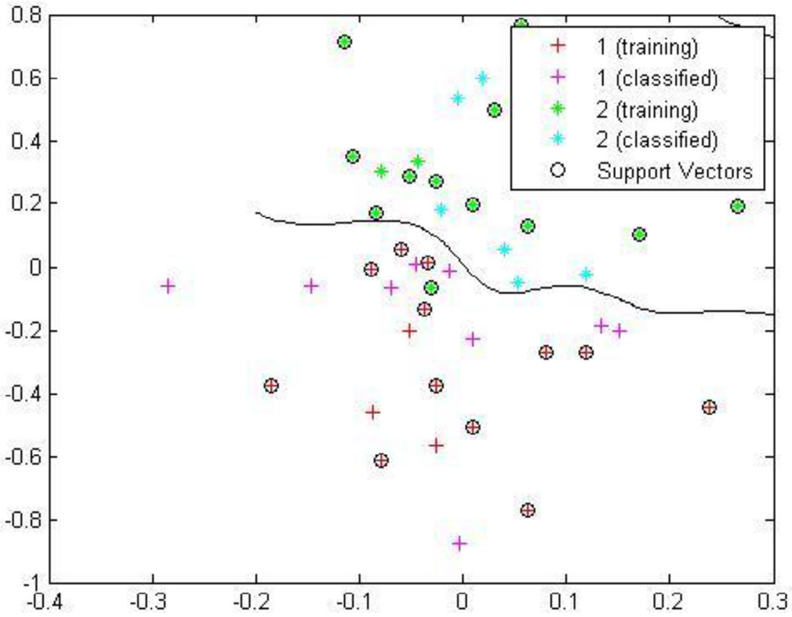


Fig. 5 – Schematic view of SVM performance.

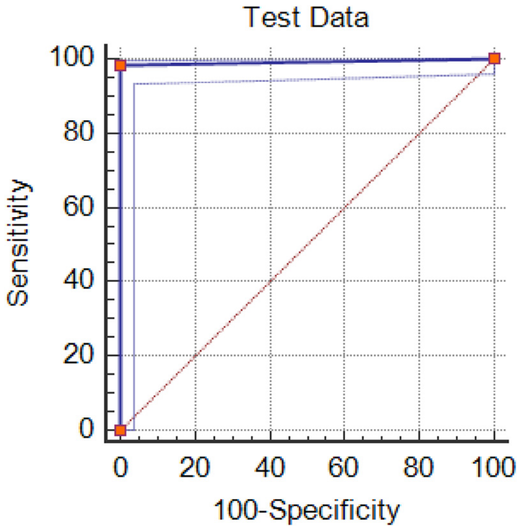


Fig. 6 – ROC plot of SVM performance.

number of normal cases identified as normal, FN (False Negative) is the number of CAD cases incorrectly identified as normal, TP (True Positive) is the number of CAD cases correctly identified as they are, and FP (False Positive) is the number of normal cases incorrectly identified as CAD. Sensitivity is the probability that a classifier will respond positive when used on the CAD subjects (TP/(TP + FN)). Specificity is the probability that a classifier will respond negative when used on the normal subjects (TN/(TN + FP)), and finally, accuracy is the ratio of the number of correctly classified data to the total number of data ((TP + TN)/(TP + FN + TN + FP)).

Classification was performed thrice using components obtained through PCA as inputs. Table 4 provides a summary of the classification results. It can be observed that PCA combined with optimized SVM classifier yields the accuracy of 99/21%, sensitivity of 98/43%, and specificity of 100%.

The diagnostic performance of a test or the accuracy of a test to discriminate diseased cases from normal cases is evaluated using Receiver Operating Characteristic (ROC) curve analysis. This plot shows a good accuracy if the left line is near vertical axis. Considering Fig. 6, it can be interpreted that the proposed SVM performs in a right way.

6. Discussions

In this study the data have been used from MIT-BIH data bank. Considering that most of the recent articles have cited this databank as their reference for the studied data, comparison between results has been conducted efficiently in Table 5.

In Karimi et al. [38], Wavelet Packet Transform (WPT) and DWT are conducted on Heart Sound Signals (HSS) in order to classify CAD and Normal subjects which has led to the accuracy of 85%. In Lee et al. [39], nonlinear and linear features have been extracted and fed into SVM classifier. The accuracy of this method is 90%. Kim has applied multiple Discriminant Analyses with linear and nonlinear features using different classifiers on HRV and has achieved the accuracy of 84.6. Zhao and Ma [41] used Teager Energy Operator (TEO) and Empirical Mode Decomposition (EMD) to diagnose CAD with HSS with the accuracy of 80%. Lee studied the relation of carotid arterial wall thickness and CAD and reached the accuracy of 90% using SVM as a classifier. According to Babaoğlu et al. [15], the data have been selected from stress test using Binary Particle Swarm Optimization (BPSO) and Genetic Algorithm (GA) and are used for classification with accuracy of 81.46% using SVM classifier. In another study [43], the number of features has been decreased using Principal Component Analysis (PCA) and the accuracy has been 79.17%. Giri et al. [4] have applied DWT to decompose the HRV signals. Independent Component Analysis (ICA), PCA, and Linear Discriminant Analysis (LDA) have reduced the number of features and the more accuracy from this study is 96.8% using Gaussian Mixture Model.

The extracted ECGs from two classes of subjects (normal and CAD) have been influenced by morphological changes that make it difficult to identify P-QRS-T waves; so, the unprocessed ECG signal does not provide the proper information that is required. Features that are extracted from this signal in time, frequency and nonlinear domains are presented as a set of quantifiable properties. The samples in a feature space are scattered in a way which leads to the misclassification and significant errors. Additionally, in some directions, features in both classes may have the same behavior in n dimensional space. It means that solving the problem with redundant dimensions has more computational complexity and takes more time. Selecting a classifier which classifies unseen data with high accuracy was one

Table 5 – Comparison of recent studies for diagnosis of CAD.

Authors	Techniques/features used	Classifiers	Input data	Accuracy
Karimi et al. [38]	Wavelet analysis	Neural network	Heart sound signal	85
Lee et al. [39]	Linear and nonlinear parameters	SVM classifier	HRV	90
Kim et al. [40]	Multiple discriminant analysis with linear and nonlinear features	Different classifiers	HRV	72/5–84/6
Zhao and Ma [41]	Empirical mode decomposition Teager Energy Operator	Back propagation neural network	ECG	80
Lee et al. [42]	HRV, carotid arterial wall thickness, CPAR	SVM	HRV	85–90
Babaoğlu et al. [43]	Binary particle swarm optimization	SVM	Doppler signal	81.46
Babaoğlu et al. [15]	PCA for dimension reduction	SVM		79.17
Giri et al. [4]	HRV signals and ICA	GMM	HRV	96/8
In this work	HRV signals and PCA and fuzzy entropy	SVM	HRVs	99.2

of the most important parts of the study. Since classification algorithms that map input data to a category, mostly, are based on mathematical functions, the performance of classifiers is dependent on the value of parameters that define these functions. This underlines the necessity of meticulous parameter optimization of the SVM algorithm in this paper.

According to the ECG diagnostic problem, HRV signal is used as a basic signal, which is not influenced by subtle changes. Considering features as a set of correlated variables, PCA is a statistical approach that extracts a set of principal components which are linearly uncorrelated variables. These principal components were used as a new set of features with lower number of features. This transformation was in order to extract features with the largest possible variance with guaranteed independency. Using PCA, the most significant features that discriminate the subjects with CAD and normal subjects have been extracted and the issue of overlapping features and redundant dimensions was solved. The classifier used in this study was support vector machine (SVM) which is a non-probabilistic binary linear classifier. Based on training dataset, SVM algorithm builds a model that classifies data into two classes. The classification of SVM can be nonlinear using Kernel function. The performance of SVM is dependent on the variable parameters of this function. A classification using traditional SVM algorithm led to the accuracy of 90/62%, sensitivity of 87/5% and specificity of 93/75%, although utilization of the proposed algorithm to indicate the best values for the parameters of SVM kernel function, the accuracy of 99/21%, sensitivity of 98/43% and specificity of 100% has been achieved; in other words, the optimized SVM offers an insurmountable automated diagnostic accuracy due to the high performance and low computational costs. This method is a promising candidate for clinical applications and can be used in routine diagnostic protocols in cardiac diagnostic centers in future.

The major contribution of our work considering other existing works is that, the proposed method uses less number of features to obtain the highest accuracy of 99.2%. In this study the number of data that have been used for training is 10, which makes it a robust solution for diagnosis of the cases in which the number of available data is small. Therefore, the proposed system in this paper is suitable for clinicians and can be installed in the hospitals to detect CAD automatically using HR signals. As we are using 11 features, the diagnosis of CAD will be fast. Mobile systems can be used to send heart signal and check the situation of the patients online. In this work CAD can be diagnosed in early stage. This has been done in the process of feature selection with the vision that small changes must not be ignored. The parameters of SVM have been optimized to consider these small changes. Therefore, the proposed system can help save the life of the CAD patients.

REFERENCES

[1] N.D. Wong, Epidemiological studies of CHD and the evolution of preventive cardiology, Nat. Rev. Cardiol. 11 (2014) 276–289.

[2] American Heart Association, Coronary artery disease, 2016. http://www.heart.org. (Accessed 10 August 2016).

[3] National Heart, Lung and Blood Institute, What is coronary heart disease?, 2015. http://www.nhlbi.nih.gov/. (Accessed 25 November 2015).

[4] D. Giri, U.R. Acharya, R.J. Martis, S.V. Sree, T.-C. Lim, T. Ahmed VI, et al., Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform, Knowl. Based Syst. 37 (2013) 274–282.

[5] U.R. Acharya, K.P. Joseph, N. Kannathal, S. Suri, Heart rate variability: a review, Med. Biol. Eng. Comput. 44 (12) (2006) 1031–1051.

[6] U.R. Acharya, N. Kannathal, S.M. Krishnan, Comprehensive Analysis of cardiac health using heart rate signals, Physiol. Meas. 25 (2004) 1139–1151.

[7] F. Lombardi, Chaos theory, heart rate variability, and arrhythmic mortality, Circulation 101 (1) (2000) 8–10.

[8] Task Force of the European Society of Cardiology and North American Society of Pacing and Electrophysiology, Heart rate variability: standards of measurement, physiological interpretation and clinical use, Eur. Heart J. 17 (1996) 354–381.

[9] Y. Isler, M. Kuntalp, Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure, Comput. Biol. Med. 37 (2007) 1502–1510.

[10] A. Schumann, N. Wessel, A. Schirdewan, K.J. Osterziel, A. Voss, Potential of feature selection methods in heart rate variability analysis for the classification of different cardiovascular diseases, Stat. Med. 21 (15) (2002) 2225–2242.

[11] A.R. Gujjar, T.N. Sathyaprabha, D. Nagaraja, K. Thennarasu, N. Pradhan, Heart rate variability and outcome in acute severe stroke: role of power spectral analysis, Neurocrit. Care 1 (3) (2004) 347–354.

[12] R.M. Carney, K.E. Freedland, P.K. Stein, J.A. Skala, P. Hoffman, A.S. Jaffe, Change in heart rate and heart rate variability during treatment of depression in patients with coronary artery disease, Psychosom. Med. 62 (2000) 639–647.

[13] U.R. Acharya, O. Faust, V. Sree, G. Swapna, R.J. Martis, N.A. Kadri, et al., Linear and nonlinear analysis of normal and CAD-affected heart rate signals, Comput. Methods Programs Biomed. 113 (2014) 55–68.

[14] J. Shlens, A tutorial on principal component analysis, 25 March 2003.

[15] I. Babaoğlu, O. Findik, M. Bayrak, Effects of principle component analysis on assessment of coronary artery diseases using support vector machine, Expert Syst. Appl. 37 (3) (2010) 2182–2185.

[16] C. Corts, V. Vapnik, Support vector networks, Mach. Learn. 20 (1995) 273–297.

[17] V. Vapnik, Statistical Learning Theory, Springer, Berlin, Germany, 1998.

[18] E. Osuna, R. Freund, F. Girosit, Training support vector machines: an application to face detection, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Puerto Rico, pp. 130–136, 1997.

[19] Y.H. Liu, Y.T. Chen, Face recognition using total margin-based adaptive fuzzy support vector machines, IEEE Trans. Neural Netw. 18 (1) (2007) 178–192.

[20] J.C. Burges, A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discov. 2 (1998) 121–167.

[21] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P. Ch. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex Physiologic Signals, Circulation 101 (23) (2000) e215–e220. http://circ.ahajournals.org/content/101/23/e215.full. (Accessed 13 June 2016).

[22] Normal Sinus Rhythm RR Interval Database.

[23] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Trans. Biomed. Eng. 32 (3) (1985) 230–236.

[24] S.L. Marple, Digital Spectral Analysis, Prentice-Hall International, 1987.

[25] M.P. Tarvainen, J.-P. Niskanen, J.A. Lipponen, P.O. Ranta-Aho, P.A. Karjalainen, Kubios HRV—heart rate variability analysis software, Comput. Methods Programs Biomed. 113 (2014) 210–220.

[26] M. Brennan, M. Palaniswami, P. Kamen, Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability, IEEE Trans. Biomed. Eng. 48 (11) (2001) 1342–1347.

[27] J.P. Zbilut, N. Thomasson, C.L. Webber, Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals, Med. Eng. Phys. 24 (2002) 53–60.

[28] Y. Fusheng, H. Bo, T. Qingyu, Approximate entropy and its application in biosignal analysis, chapter 3, in: M. Akay (Ed.), Nonlinear Biomedical Signal Processing: Dynamic Analysis and Modeling, vol. II, IEEE Press, New York, 2001, pp. 72–91.

[29] J.A. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol. 278 (2000) H2039–H2049.

[30] N.V. Thakor, S. Tong, Advances in quantitative electroencephalogram analysis methods, Ann. Rev. 6 (2004) 453–495.

[31] C.K. Peng, S. Havlin, J.M. Hausdorf, J.E. Mietus, H.E. Stanley, A.L. Goldberger, Fractal mechanisms and heart rate dynamics, J. Electrocardiol. 8 (Suppl.) (1996) 59–64.

[32] C.-K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, Chaos 5 (1995) 82–87.

[33] P. Grassberger, I. Procassia, Measuring the strangeness of strange attractors, Physica D 9 (1983) 189–208.

[34] U.R. Acharya, S.V. Sree, S. Chattopadhyay, W. Yu, A.P.C. Alvin, Application of recurrence quantification analysis for the automatic identification of epileptic EEG signals, Int. J. Neural Syst. 21 (3) (2011) 199–211.

[35] V. David, A. Sanchez, Advanced support vector machines and kernel methods, Neurocomputing 55 (2003) 5–20.

[36] J. Ren, ANN vs. SVM: which one performs better in classification of MCCs in mammogram imaging, Knowl. Based Syst. 26 (2012) 144–153 (in press).

[37] P. Ping-Feng, M.-F. Hsu, M.-C. Wang, A support vector machine based model for detecting top management fraud, Knowl. Based Syst. 24 (2011) 314–321.

[38] M. Karimi, R. Amirfattahi, S. Sadri, S.A. Marvasti, Noninvasive detection and classification of coronary artery occlusions using wavelet analysis of heart sounds with neural networks. In The 3rd IEE International Seminar on Medical Applications of Signal Processing (pp. 117–120), 2005.

[39] H.G. Lee, K.Y. Noh, K.H. Ryu, Mining biosignal data: coronary artery disease diagnosis using linear and nonlinear features of HRV, in: T. Washio, Z.-H. Zhou, J.Z. Huang, X. Hu, J. Li, C. Xie, et al. (Eds.), Emerging Technologies in Knowledge Discovery and Data Mining, vol. 4819, Springer Berlin Heidelberg, 2007, pp. 218–228. of Lecture Notes in Computer Science.

[40] W.S. Kim, S.H. Jin, Y.K. Park, H.M. Choi, A study on development of multi-parametric measure of heart rate variability diagnosing cardiovascular disease, World Congress on Medical Physics and Biomedical Engineering 2006, Volume 14 of the series IFMBE Proceedings pp 3480–3483.

[41] Z. Zhao, C. Ma, An intelligent system for noninvasive diagnosis of coronary artery disease with EMD-TEO and BP Neural Network. In International Workshop on Education Technology and Training and International Workshop on Geoscience and Remote Sensing. (pp. 631–635).volume 2, 2008.

[42] H.G. Lee, A data mining approach for coronary heart disease prediction using HRV features and carotid arterial wall thickness; BioMedical Engineering and Informatics, 2008.

[43] I. Babaoğlu, O. Findik, E. Ulker, A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine, Expert Syst. Appl. 37 (4) (2010) 3177–3183.